

## **METHOD OF MAKING HIGHLY DISCRIMINATING OPTICAL EDGE FILTERS AND RESULTING PRODUCTS**

### 5        **Cross Reference To Related Applications**

This application claims the benefit of United States Provisional Patent Application Serial Number 60/468,245 filed by Ligang Wang on May 6, 2003 and entitled "Highly Discriminating Optical Edge Filters", which application is incorporated herein by reference.

### **Field of Invention**

10        This invention relates to methods of making optical edge filters and to the resulting improved filters. The filters provide a factor of two improvement in edge steepness and a factor of two reduction in transmission loss.

### **Background of the Invention**

#### **A. Optical edge Filters and their use**

15        Optical edge filters are key components in important systems for optical measurement and analysis including Raman spectroscopy and fluorescence microscopy. Optical edge filters are used in these systems to block unwanted light that would otherwise generate spurious optical signals and swamp the signals to be detected.

20        Optical edge filters block unwanted light having wavelengths above or, alternatively, below a chosen "transition" wavelength  $\lambda_T$  while transmitting light on the unblocked side of  $\lambda_T$ . Edge filters which transmit optical wavelengths longer than  $\lambda_T$  are called long-wave-pass filters (LWP filters), and edge filters which transmit wavelengths shorter than  $\lambda_T$  are short-wave-pass or SWP filters.

25        Referring to the drawings, Figs. 1A and 1B schematically illustrate the spectral transmission of idealized long-wave-pass and short-wave-pass filters respectively. As can be seen from Fig. 1A, a LWP filter blocks light with wavelengths below  $\lambda_T$  and transmits light with wavelengths above  $\lambda_T$ . As shown in Fig. 1B, a SWP filter transmits light with wavelengths

below  $\lambda_T$  and blocks light with wavelengths above  $\lambda_T$ .  $\lambda_T$  is the wavelength at which the filter “transitions” from blocking to transmission, or vice versa.

While an ideal edge filter has a precise transition wavelength  $\lambda_T$  represented by a vertical line at  $\lambda_T$ , real edge filters change from blocking to transmission over a small range of wavelengths and are more accurately represented by a non-vertical but steeply sloped line near  $\lambda_T$ . Similarly, while an ideal edge filter transmits all light in the transmission region (transmission  $T=1$ ), real filters invariably block a small portion of the light to be transmitted ( $T<1$ ). The steepness of the line and the proportion of the light transmitted are important parameters in many applications.

Edge filters are particularly useful in optical measurement and analysis systems that use light of one wavelength to excite a sample and measure or view an optical response of the excited sample at other wavelengths. The excitation light is delivered to the sample by an excitation light path, and the optical response of the sample is delivered to the eye or measuring instrument by a collection path. Edge filters can be used to block spurious light from the excitation path and to block excitation light from entry into the collection path. The steeper the filter edge, the more effectively spurious signals are blocked. The lower the transmission loss, the more light from the sample reaches the measuring instrument.

Raman spectroscopy is one such optical analysis system. It is based on the fact that when molecular material is irradiated with high intensity light of a given wavelength  $\lambda$ , a small portion of the incident light scattered by the material will be shifted in wavelength above and below  $\lambda$ . This Raman shifting is attributed to the interaction of the light with resonant molecular structures within the material, and the spectral distribution of the Raman-shifted light provides a spectral “fingerprint” characteristic of the composition of the material. As a practical example, one can use a Raman probe to identify the contents of a bottle without opening the bottle.

Fig. 2 is a simplified schematic diagram of a Raman probe designed to excite and collect the long wavelength portion of Raman-shifted light from a sample. In essence, the probe comprises an optical fiber excitation path, and a fiber collection path. Edge filters are disposed in the respective paths.

In operation, excitation light from a laser 24 passes through the fiber path 22 and edge filter 22A to illuminate a portion of the sample 21 with high intensity light of a wavelength  $\lambda$ . Light scattered from the sample 21 passes through edge filter 23A and then through fiber collection path 23 to a spectral analyzer 25 where the “fingerprint” of the sample is determined.

5            Since the fiber 22 through which the excitation signal passes is composed of molecular material, a small portion of the excitation light will be shifted in wavelength by the Raman effect in the fiber. This shifted light must be eliminated to prevent false readings. The removal can be accomplished by disposing a SWP edge filter 22A between the fiber 22 and sample 21. SWP edge filter 22A, having a transition wavelength just above the laser wavelength, blocks both long  
10            wavelength Raman scattering from the fiber and long wavelength noise from the laser.

             However, in some cases, a filter 22A is not needed. For instance, if the excitation signal is direct laser radiation that does not pass through molecular material at any appreciable length, the filter 22A is not necessary. An example would be direct laser radiation illuminating a sample 21 through a vacuum.

15            The light scattered from the sample 21 is a mixture of unshifted scattered excitation light (Rayleigh scattering) and Raman-shifted light. The scattered excitation light would not only swamp the analyzer, it would also excite spurious Raman scattering in the collection fiber. Thus the unshifted excitation light should be removed from the collection path. This can be accomplished by disposing a long pass edge filter 23A between the sample 21 and the collection  
20            fiber 23, the long pass filter having a transition wavelength  $\lambda_T$  just below the excitation wavelength  $\lambda$ . This arrangement ensures that the light reaching the analyzer is predominantly the long wavelength Raman-shifted light from the sample. Analogous arrangements using edge filters can be used to analyze short wavelength Raman-shifted light.

             Edge filters are equally useful in fluorescence microscopy. Here excitation light is used  
25            to excite longer wavelength emission from fluorescent markers. The markers can be fluorescent atoms chemically bonded to a biological molecule to track the molecule in a body or cell. Edge filters are used, as in Raman spectroscopy, to reject spurious low wavelength light from the excitation path and to reject excitation light from the collection path.

It should now be clear that the steeper the filter slope at the transition wavelength  $\lambda_T$  the greater the amount of spurious light that can be filtered out. In addition, the steeper the slope, the greater the amount of shifted light from the sample that will reach the analyzer. Similarly, higher levels of transmission of the shifted light through the filters provide more light for analysis. Higher edge filter blocking provides better rejection of the laser excitation light from the spectrum analyzer, thus decreasing the noise and improving both specificity and sensitivity of the measurement. Higher edge-filter transmission enables the maximum signal to reach the analyzer, further improving the signal-to-noise ratio and hence the measurement or image fidelity. A steeper filter edge also permits shifts to be resolved much closer to the excitation wavelength, thus increasing the amount of information from the measurement.

### **B. Edge Filter Structure and Conventional Fabrication**

Fig. 3 is a simplified schematic illustration of an optical edge filter 30 comprising a transparent substrate 31 having a flat major surface 32 supporting many thin coatings 33A, 33B. The thickness of the coatings is exaggerated and the number is reduced for purposes of illustration. Coatings 33A and 33B are typically alternating and of different respective materials chosen to present markedly different indices of refraction (index contrast). The coating indices and thicknesses are chosen and dimensioned to filter impinging light by interference effects in a desired manner. Specifically, if a light beam 34 impinges on the filter, a first wavelength portion 34T of the beam is transmitted and a second wavelength portion 34R is reflected and thus rejected by the filter. What is transmitted and what is reflected depends on the precise thickness and indices of the thin coatings. There are two basic types of thin-film edge filters: those based on “soft coatings” and those based on “hard coatings,” both of which are typically manufactured by an evaporation technique (either thermal evaporation or electron-beam evaporation). Hard coating filters, however, may also be manufactured by non-evaporative techniques such as ion-beam sputtering.

Soft coatings imply literally what the name suggests—they are physically soft and can be readily scratched or damaged. They are fairly porous, which also means they tend to be hygroscopic (absorb water vapor) leading to dynamic changes in the film index and hence the resulting filter spectrum in correlation to local humidity. There are two main reasons soft

coatings are used. First, an advantageous larger index contrast can be realized with soft coatings. (The index contrast is the relative difference between the index of refraction of the low-index material and that of the high-index material.) For example, many high-performance soft-coated filters are made using sodium aluminum fluoride (“cryolite”), with a chemical composition of  $\text{Na}_3\text{AlF}_6$  and an index of about 1.35 for visible wavelengths, and zinc sulfide, with a chemical composition of  $\text{ZnS}$  and an index of about 2.35. The second reason for using these materials is that the evaporation process can be controlled well for these materials, largely because they have relatively low melting temperatures. Hence it is possible to maintain fairly accurate control over the layer thicknesses even for filter structures with many 10’s of layers and perhaps even up to 100 layers. As described above, edge filter performance is measured by edge steepness, depth of blocking, and high transmission with low ripple. A larger index contrast and a larger number of layers both yield more steepness and more blocking. High transmission with low ripple is improved with more layers and higher layer thickness accuracy. For these reasons the highest performance conventional thin-film edge filters have been made with soft-coating technology.

Hard coatings are made with tougher materials (generally oxides), and result from “energetic” deposition processes, in which energy is explicitly supplied to the film itself during the deposition process. This is accomplished with a beam of ions impinging directly on the coating surface. The ion bombardment acts to “hammer” the atoms into place in a more dense, less porous film structure. Such processes are usually called ion-assisted deposition (IAD) processes. High-performance edge filters have been made with ion-assisted electron-beam evaporation. Typically the index contrast available with hard-coating (oxide) thin-film materials is not as high as that of the soft-coating materials, and consequently more layers must be deposited to achieve a comparable level of performance. This problem, coupled with the more difficult to control deposition rates and overall processes of high-melting-temperature oxides, leads to much more stringent requirements on the layer-thickness control techniques to achieve a reasonable level of layer thickness accuracy for good edge steepness and high, low-ripple transmission.

For the best edge filters, some kind of “optical monitoring” (direct measurement of filter transmission or reflection during deposition) is necessary to determine when to terminate the

deposition of each layer. Optical monitoring can be performed on the actual filters of interest or on “witness pieces” often positioned in the center of the deposition chamber. There are three basic types of optical monitoring algorithms. The first is often called “drop-chip” monitoring, and is based on measuring the transmission (or reflection) vs. time through a new witness piece  
5 for each new layer. Since the theoretical transmission vs. time can be calculated accurately for each layer deposited on a blank piece of glass, then a good comparison between the measured and theory curves can be made independent of the history of the deposition (thickness errors in previous layers). This technique is accurate and useful for layers of arbitrary thickness, but it is cumbersome, especially for filters comprised of at least many 10’s of layers.

10 The second type of monitoring is called “turning-point” monitoring, and is used for depositing layers that are precisely a quarter of a wavelength in thickness (or multiples thereof). The technique is based on the fact that the transmission vs. time reaches a turning point (or extremum) at each multiple of a quarter wave of thickness, so an algorithm is developed to cut  
15 layers precisely at the turning points. The elegant feature of this method is that there is inherent compensation for layer thickness errors from previous layers, so long as one adheres to the rule of cutting exactly at turning points. It thus works extremely well even for very thick coatings with even hundreds layers (it is the basis for manufacturing very high-performance filters for DWDM telecom applications, which can have as many as 200 – 400 quarter-wave layers).

20 The third type of monitoring is called “level monitoring,” and is applicable for non-quarter-wave thick layers. Monitoring can be done through the actual filters or through witness piece(s). The concept is to cut layers at predetermined transmission levels, based on a calculated prediction of transmission vs. time for the entire structure. However, because small layer errors  
25 lead to large variations in the absolute transmission values, one must instead rely on cutting at the correct transmission level relative to the local maximum and minimum values. Hence the method works well only for non-quarter-wave thick layers that are more than a half-wave thick, so that there is both a maximum and a minimum transmission value in the transmission vs. time curve for that layer. Even in this case, this method does not contain inherent compensation for  
30 errors in the thickness of previously deposited layers, and thus is not as forgiving as the turning-point method. However, to obtain an edge filter with high transmission and low ripple requires

primarily non-quarter-wave thick layers, and hence turning-point monitoring is not applicable for edge filters.

Accordingly there is a need for an improved method of making an optical edge filter and  
5 for improved edge filters having increased edge steepness and reduced transmission.

### **Summary of the Invention**

These problems are addressed and a technical solution achieved in the art by highly discriminating optical edge filters and methods of making the same according to the present invention. In particular, the optical edge filters have an edge steepness less than about 0.8% as  
10 measured by dividing (a) the edge width from the 50% transmission wavelength to the optical density 6 ("OD6") wavelength by (b) the 50% transmission wavelength. The optical filters also have an average transmission above about 95%. The methods for making such filters accurately determine when deposition of each layer of the filter should terminate. The methods include  
15 calculating theoretical transmission data for a layer of the filter and calculating an expected deposition duration for the layer. The methods also include measuring transmission through the layer during deposition for a period less than the expected deposition duration. When the measuring period elapses, a new deposition duration is calculated based upon the theoretical transmission data and the measured transmission data, thereby providing an accurate deposition duration for the layer.

### **Brief Description of the Drawings**

The advantages, nature and various additional features of the invention will appear more fully upon consideration of the illustrative embodiments now to be described in detail in connection with the accompanying drawings. In the drawings:

25 Figs. 1A and 1B are schematic graphical illustrations showing the spectral transmission of long-wave-pass and short-wave-pass optical edge filters, respectively;

Fig. 2 is a schematic diagram of a conventional Raman probe designed to excite and collect the long wavelength portion of Raman-shifted light from a sample;

Fig. 3 is a schematic drawing of an idealized optical edge filter;

Fig. 4 is a schematic diagram of apparatus useful in making an optical edge filter in  
5 accordance with the invention;

Fig. 5 is a process flow illustrating the process of manufacturing a long-wave-pass filter in accordance with the invention; and

Fig. 6 is a process flow illustrating the process of manufacturing a short-wave-pass filter in accordance with the invention.

10 Figs. 7 and 8 are transmission and optical density spectra, respectively, of an LWP filter fabricated in accordance with the invention; and

Figs. 9 and 10 are transmission and optical density spectra, respectively, of an SWP filter fabricated in accordance with the invention.

It is to be understood that these drawings are for illustrating the concepts of the invention  
15 and, except for data graphs, are not to scale.

### **Detailed Description of the Invention**

The invention has four aspects: A) apparatus programmed to make improved optical edge filters; B) the method of making the improved filters; C) the improved filters made by the method and D) applications of the improved filters. These aspects will be described in the order  
20 presented.

#### **A. Apparatus For Making Optical Edge Filters In Accordance With the Invention**

Optical edge filters in accordance with the invention are made using a computer-controlled deposition system. Advantageously the deposition is an ion beam sputtering deposition system using a beam assist source for depositing hard coatings and having an integral  
25 optical monitoring system to monitor deposition. A data processor, responsive to signals from



the monitoring system, processes these signals and directs the growth of improved optical edge filters in accordance with algorithms described below.

Fig. 4 is a schematic view of an advantageous computer-controlled deposition apparatus 400 for practicing ion beam sputtering disposition, ion beam assisted deposition and optical monitoring under data processor control. The apparatus 400 comprises a vacuum chamber 401 having an interior accessible by a sealable port 402 such as an o-ring sealed door. The chamber 401 includes a pump port 402 for coupling to one or more evacuation pumps (not shown) such as mechanical and low pressure cryogenic pumps.

Within the chamber 401 is a substrate mount 403 which is advantageously a rotatable spindle mount. Also within the chamber are a plurality of material targets, a deposition ion beam source 404, and a mounted substrate 405. One of the targets 406 is positioned in relation to the ion beam source and the substrate so that an ion beam from the source will sputter material from the target onto the substrate in a substantially uniform layer. Typically there are separate targets for each material to be deposited on the substrate, and the targets are movable to and from the sputtering position. There are usually two material targets: one a high optical index material and the other a low index material. The targets are movable so that deposition can be switched automatically from one material to the other.

An assist ion beam source 407 is disposed in the chamber 401 in position to bombard the substrate 405 with an ion beam. During the deposition of any one material, the substrate is simultaneously bombarded by sputtered material and by ions from the assist ion beam source 407. The assist ion bombardment serves to energize the adhesion of particles onto the surface of the substrate and thereby produce a substantially more dense ("hard") film than would otherwise be formed by sputter deposition. See J.M.E. Harper *et al.*, "Modification of Thin Film Properties by Ion Bombardment During Deposition," in Ion Bombardment Modification of Surfaces, Ed. By O. Auciello and R. Kelly, from Beam Modifications of Materials, Vol. 1 (Elsevier, Amsterdam, 1984).

Both the sputter deposition and the assist bombardment can be activated by a deposition control system 408. The control system 408 can start or stop deposition precisely by turning on and off the deposition ion beam source 404 or by removing or engaging a mechanical shutter

(not shown) that covers the substrate. Further details concerning ion beam sputtering deposition systems can be found in United States Patent No. 4,142,958 issued to David Wei *et al.* on March 6, 1979 and U.S. Patent No. 4,793,908 issued to Gene Scott *et al.* on December 27, 1988, both of which are incorporated herein by reference. *See also* Juergen Becker, "Ion Beam Sputtering,"  
5 Handbook of Optical Properties, Vol. 1, Thin Films for Optical Coatings, Ed. by R.E. Hummel and K.H. Guenther (CRC Press, Boca Raton, 1995).

The optical monitoring system advantageously comprises a light source 409, filters 411A and 411B, detectors 412 and 413 and a data processor 414. The light source 409 can be any sufficiently bright, broadband light source, such as a quartz halogen bulb or a Xenon discharge  
10 lamp. The detectors can include a reference detector 412 and a signal detector 413. The detectors can be semiconductor (Si) photodiodes, photomultiplier tubes or any other sensitive, low-noise detectors capable of detecting light at the monitoring wavelength.

In the advantageous arrangement shown here, a beamsplitter 415 picks off a portion 416 of the monitoring light beam 417 from source 409 and detects it with reference detector 412.  
15 The signal detector 413 detects the portion 418 of the beam 417 that passes through the coated substrate (or a "witness" substrate) being monitored. Advantageously, the filters 411A and 411B are positioned to ensure that a sufficiently narrow band of wavelengths is monitored.

The filters 411A, 411B can be narrow band interference filters, adjustable diffraction-grating monochromators or combinations thereof. Advantageously an interference filter is used  
20 for reference filter 411A and a monochromator is used for detector filter 411B. Further details concerning optical monitoring of thin films as they are being deposited can be found in United States Patent No. 6,649,208 issued to Wayne Rodgers on November 18, 2003, which is incorporated herein by reference. *See also* H. Angus Macleod, Thin Film Optical Filters (3<sup>rd</sup> Ed., Institute of Physics, Bristol, 2001) and H.A. Macleod, "Turning Value Monitoring of Narrow-Band All-Dielectric Thin-Film Optical Filters," Optica Acta, vol. 19, pp. 1-28(1972).  
25

The data processor 414 collects data from the signal and reference detectors 413, 412, implements the mathematics associated with optical monitoring algorithms and instructs the deposition control system 408 when to stop depositing any given thin film layer based on the result of the mathematical calculations prescribed in the algorithm.

## **B. Methods of Making the Improved Edge Filters**

The manner in which the data processor 414 controls the apparatus 400 via the deposition control system 408 to generate the optical edge filters according to the present invention will now be described. In particular, the data processor 414 is programmed to instruct the apparatus 400 when to stop depositing each layer of the optical edge filter being manufactured. The data processor 414 follows two separate processes in determining when deposition of a layer should terminate depending upon whether a long-wave-pass or a short-wave-pass filter is being manufactured. These two processes will be described in turn.

Fig. 5 illustrates the process flow executed by the data processor 414 when manufacturing a long-wave-pass ("LWP") filter according to the present invention. However, prior to initiating the process of Fig. 5, a design for the LWP filter is prepared. In the exemplary embodiment, the LWP filter has N layers and comprises two materials: a low-refractive-index material and a high-refractive-index material. The exemplary initial design for a steep LWP edge filter includes a quarter-wave ("QW") stack of  $(0.5H L 0.5H)^N$ , where L and H represent layers of high and low index materials with a quarter-wave of optical thickness at the reference wavelength. The reference wavelength is chosen so that the longer-wavelength edge of the QW stopband is close to the desired transition wavelength of the LWP filter design.

Once the initial design is setup, a desired target spectrum is constructed, which typically includes the wavelength ranges of both the blocking and passband regions, as well as the required blocking level and minimum transmission and allowable ripple within the passband. The edge steepness is thus indirectly defined as the wavelength separation between the blocking region and the passband.

The layer thicknesses of the initial design are then optimized against the target spectrum by an optimization routine known in the art. Exemplary optimization routines include the variable-metric or simplex methods implemented in standard commercial thin-film design software packages, such as TFCalc by Software Spectra, Inc., and The Essential Macleod by Thin Film Center Inc. Usually, with the proper choice of the initial design, the optimization quickly converges and the optimized structure is not very different from the initial structure. Special treatments are required for the first layer (toward substrate) and the last layer. The

optical thickness of the first layer is manually increased by at least two QWs. As for the last layer, it sometimes becomes too thin and, consequently, should be eliminated. The structure should be re-optimized whenever there is any modification to the layer thickness.

With the LWP filter design at hand, the data processor 414 receives design data and deposition rate data as input at 501. The design data describes the designed thin-film structure of the LWP filter with a physical thickness  $d_i$  and an index  $n_i$  for each  $i^{\text{th}}$  layer.  $n_i$  is either  $n_L$  or  $n_H$ , where  $n_L$  is the refractive index of the low-index material and  $n_H$  is the refractive index of the high-index material.  $n_L$  and  $n_H$  are each known as a function of wavelength  $\lambda$ . The deposition rate data describes the known starting deposition rate of the deposition apparatus 400 shown in Fig. 4 for each of the two materials. In the exemplary embodiment, the deposition rate data is within about +/- 5% of the actual deposition rate and is in units of Å/sec. The starting rate estimates for each of the two materials are referred to as  $r_L$  and  $r_H$ , and hence each layer will have a starting rate estimate  $r_i$  depending on whether it is made of low-index or high-index material.

At 502, the transmission  $T_i$  as a function of physical thickness  $d$  for each  $i^{\text{th}}$  layer is calculated at a series of wavelengths in the transmissive band of the finished filter. Consequently, the calculations at 502 result in a series of curves  $T_i$  vs.  $d$  at each of the series of wavelengths in the transmissive band of the finished filter. Such calculations are performed using standard mathematical methods for calculating the optical properties of thin-film filters. See, for example, H. A. Macleod, Thin-film Optical Filters, 3<sup>rd</sup> edition (Institute of Physics, Bristol, 2001).

Advantageously, instead of calculating each  $T_i$  vs.  $d$  curve at one corresponding wavelength in the series of wavelengths, each  $T_i$  vs.  $d$  curve is calculated by averaging a plurality of  $T_i$  vs.  $d$  curves calculated at a range of wavelengths surrounding the corresponding wavelength. For instance, assume that the series of wavelengths includes 501 nm and 502 nm. Instead of calculating a  $T_i$  vs.  $d$  curve at just 501 nm, this  $T_i$  vs.  $d$  curve is advantageously calculated by averaging  $T_i$  vs.  $d$  curves calculated at, for instance, 500 nm, 501 nm, and 502 nm. Further, the  $T_i$  vs.  $d$  curve at 502 nm is advantageously calculated by averaging  $T_i$  vs.  $d$  curves

calculated at, for instance, 501 nm, 502 nm, and 503 nm. One skilled in the art will appreciate that invention is not limited to this averaging procedure and the range of wavelengths used.

At 503, an optical monitoring wavelength  $\lambda_m$  is selected from the series of wavelengths in the transmissive band of the finished filter, thereby identifying a single curve  $T_i$  vs.  $d$  at  $\lambda_m$  from the series of curves computed at 502. The monitoring wavelength  $\lambda_m$  is determined based on the contrast of the monitoring signal within each layer. The contrast is defined as the relative range of the monitoring signal within the layer of interest. The higher the contrast, the more robust the process flow of Fig. 5 will be with respect to random signal noise.

Advantageously, the relative separation between the monitoring and cutoff wavelengths should be at least about 2%. Once above 2%, the monitoring wavelength  $\lambda_m$  should be chosen to maximize the contrast of each layer. In addition, the optical thickness of the first layer toward the substrate has a significant impact on the signal contrast for the rest of the. Therefore, the optical thickness of the first layer should be close to three quarter-wavelengths at the monitoring wavelength  $\lambda_m$ .

Having determined the monitoring wavelength  $\lambda_m$  at 503, processing advances to 504 where it is determined how the deposition duration for each layer will be calculated. For layers that are predicted to have little error between the designed thickness  $d$  and a simulated actual thickness, deposition duration is controlled by optically monitoring transmission levels  $T_m$  through the layer during deposition. For the other layers, their deposition durations are controlled using an expected deposition time  $t_i$  based upon designed (“theoretical”) thickness  $d_i$  and deposition rate  $r_i$ . Accordingly, at 504, the data processor 414 determines which layers are to be optically monitored and which layers are to be timed using an expected deposition time.

To determine which layers will be optically monitored, the data processor 414 enters a simulation mode to simulate deposition of each of the layers of the optical filter. Only the layers that are determined by the simulation to have the least amount of error are selected for optical monitoring. The simulation mode is nearly identical to the process described below with reference to 505 to 522, except that layers are not actually deposited at 505 and 506, the processing described at 508 is skipped, and instead of actually measuring transmission data  $T_m$

vs.  $t$  at 511, it is generated.  $T_m$  vs.  $t$  is generated by adding random noise to the theoretical data  $T_i$  vs.  $d$  at  $\lambda_m$  from 502 and 503. In the exemplary embodiment, 0.2% peak-to-peak random noise is used, and the maximum amount of error (“threshold”) to select a layer for optical monitoring is to have no more than about 0.2% error from the theoretical thickness  $d_i$ . The error calculation, in this regard, is discussed in more detail below with reference to 518 and 519. The layers that are simulated to exceed the threshold amount of error are flagged to have their deposition duration controlled by the best estimate of the deposition rate  $r_i$  for that layer or from an average of the rates of the previous layers of like material (typically 10 to 20 of such layers).

After determining which layers are to be optically monitored at 504, actual deposition of the layers of the LWP filter begins at 505. In particular, the substrate is loaded into the deposition apparatus 400 of Fig. 4, the apparatus is pumped down to a vacuum, and deposition of the first layer (current layer  $i$ ) is initiated at 506. The expected deposition time  $t_i$  for layer  $i$  is calculated as the desired thickness  $d_i$  divided by the estimated deposition rate  $r_i$  for the layer or from an average of the rates of the previous layers of like material at 507. It should be noted, however, that calculation of the expected deposition time  $t_i$  at 507 may be calculated prior to beginning actual deposition of the current layer  $i$  at 506.

After calculating the expected deposition time  $t_i$ , it is determined whether the current layer  $i$  has been identified as a layer to be optically monitored for controlled deposition duration. If the current layer has not been so identified, deposition of the current layer terminates when the expected deposition time  $t_i$  from 507 expires, as shown at 509. After the expected deposition time  $t_i$  has expired, processing advances to 521 where the next layer is queued up for deposition, as shown at 510.

If it is determined at 508 that the current layer  $i$  is to be optically monitored, the actual transmission  $T_m$  is measured at 511 as a function of actual time transpired  $t$  until about 95% of the expected deposition time  $t_i$  has elapsed. Once about 95% of  $t_i$  has elapsed, a new deposition duration is calculated at 512-520. In particular, at 512, 513, and 514, using the  $T_i$  vs.  $d$  at  $\lambda_m$  curve from 502 and 503, a two-dimensional (2D) array of additional curves is generated by plotting the values of  $T_i$  against a 2D array of time vectors  $t_{jk}$ . In particular, at 513, transmission  $T_{ij}$  is generated by plotting  $T_i$  against the values  $t_j = d/r_j$ , where  $r_j$  represents, for each  $j$ , a

deposition rate having a value close to the predicted value  $r_i$ . That is, the set of all  $r_j$  values is a range of values surrounding the predicted value  $r_i$ . Accordingly,  $j$  is an index that counts the number of  $r$  values that come from the range surrounding the predicted value  $r_i$ . At 514, for each value of  $j$ , transmission  $T_{ijk}$  is generated by plotting  $T_{ij}$  against the values  $t_{jk} = t_j + \Delta t_k$ , where  $\Delta t_k$  represents various values used for a uniform time shift.

At 515, it is determined whether there is more than one extremum in the curve  $T_i$  vs.  $d$  at  $\lambda_m$ . If there is more than one extremum at 515, then each of the 2D array of curves generated at 512-514, is scaled in two ways at 516. First, the mid-point between the two extrema for each  $T_{ijk}$  curve is scaled by a factor so that it equals the mid-point between the two extrema of the measured data  $T_m$  vs.  $t$ . Second, the maximum and minimum values on each  $T_{ijk}$  curve are scaled by scaling uniformly about their mean so that the difference between the maximum and minimum for each curve  $T_{ijk}$  is the same as that on the measured curve  $T_m$  vs.  $t$ . If there is one or no extremum at 515, then the mean of each  $T_{ijk}$  curve is scaled at 517 by a uniform factor so that it is equal to the mean of the measured curve  $T_m$  vs.  $t$ .

After scaling at 516 or 517, processing advances to 518 where error is calculated. For each of the 2D array of generated  $T_{ijk}$  curves, the root-mean-square (RMS) error between each  $T_{ijk}$  curve and the measured curve  $T_m$  vs.  $t$  is computed. Typically this computation is performed only for data between about 10% and 95% of the expected deposition time  $t_i$ . Afterwards, the values of  $j$  and  $k$  that yield a minimum RMS error at 518 are identified at 519. Therefore, the curve  $T_{ijk}$  vs.  $t_{jk}$  is taken to be the best approximation of the actual curve  $T_m$  vs.  $t$  for layer  $i$ . At 520, the curve  $T_{ijk}$  vs.  $t_{jk}$  is compared against the design curve  $T_i$  vs.  $d$  at  $\lambda_m$  from 502 and 503, and the time  $t_{jk}$  at which the layer should be terminated is computed. When the measured time  $t$  reaches  $t_{jk}$ , the deposition for the optically monitored layer  $i$  is terminated.

After deposition of the current layer  $i$  has completed, the apparatus 400 of Fig. 4 is reconfigured at 521 to start depositing the material associated with layer  $i + 1$ , and the process loops back to 506. However, if all layers of the filter have been deposited, the LWP filter is complete, and processing stops at 522.

Turning now to Fig. 6, the process flow executed by the data processor 414 when manufacturing a short-wave-pass ("SWP") filter according to the exemplary embodiment of the

present invention is described. The process illustrated with Fig. 6 is based on the inventor's observation that the inverse of the monitoring trace (*i.e.*, the transmission of the monitoring light as a function of time) within each layer follows a sinusoidal pattern. As will be shown, only four parameters,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ , and  $\theta_{m+1}$ , are needed to fully describe the behavior of the monitoring trace for each layer.

Assume that at the monitoring wavelength the incident medium, deposited materials, and substrate are all lossless (practically true). At normal incidence, each layer may be described by its characteristic matrix:

$$\begin{bmatrix} \cos \theta & -i \sin \theta / n \\ -in \sin \theta & \cos \theta \end{bmatrix} \quad (1)$$

where  $n$  is the refractive index of the layer,  $\theta$  is the accumulated phase within the layer, and  $i$  in equations 1 through 8 (and only in these equations) is the square root of  $-1$ , or  $i = \sqrt{-1}$ . (It should be noted that although this specification is described in the context of light striking a filter at normal incidence, one skilled in the art will appreciate that this invention also applies to polarized light striking a filter at non-normal incidence.)  $\theta$  is expressed as:

$$\theta = \frac{2\pi}{\lambda_m} nd \quad (2)$$

where  $d$  is the metric thickness of the layer and  $\lambda_m$  is the monitoring wavelength. Assuming an assembly of  $m$  layers have already been deposited, the transmission of the monitoring light within the  $(m+1)^{st}$  layer, the present layer, may be expressed as:

$$T = \frac{4n_a n_s}{H} \quad (3)$$

where  $T$  is the level of monitoring signal within the present layer (transmission), and  $n_a$  and  $n_s$  are the refractive indices of the incident medium and substrate, respectively.  $H$  is defined as:



$$H = \beta_1 + \beta_2 \cos 2\theta_{m+1} + \beta_3 \sin 2\theta_{m+1} \quad (4)$$

$\beta_1$ ,  $\beta_2$ , and  $\beta_3$  are defined as:

$$\beta_1 = \frac{n_a^2 + n_{m+1}^2}{2} \left( |p|^2 + \frac{|q|^2}{n_{m+1}^2} \right) + 2n_a \operatorname{Re}(pq^*) \quad (5)$$

5 where  $\operatorname{Re}(pq^*)$  is the Real Part of the product of the complex number  $p$  and the complex conjugate of the complex number  $q$ .  $p$  and  $q$  are defined in equation (8) below.

$$\beta_2 = \frac{n_a^2 - n_{m+1}^2}{2} \left( |p|^2 - \frac{|q|^2}{n_{m+1}^2} \right) \quad (6)$$

$$\beta_3 = \left( \frac{n_a^2}{n_{m+1}} - n_{m+1} \right) \operatorname{Im}(p^* q) \quad (7)$$

10 where  $\theta_l$  is accumulated phased in the  $l_{th}$  layer.  $\operatorname{Im}(pq^*)$  is the Imaginary Part of the product of the complex conjugate of the number  $p$  and the complex number  $q$ .  $p$  and  $q$  are defined as:

$$\begin{bmatrix} p \\ q \end{bmatrix} = \prod_{l=1}^m \begin{bmatrix} \cos \theta_l & -i \sin \theta_l / n_l \\ -in_l \sin \theta_l & \cos \theta_l \end{bmatrix} \begin{bmatrix} 1 \\ n_s \end{bmatrix} \quad (8)$$

15 Based on the above equations, the relationship between the monitoring trace  $T$  and the accumulated phase  $\theta_{m+1}$  within the present layer is established. The inverse of the monitoring trace is sinusoidal, as demonstrated by the expression for  $H$ . Further, only four parameters,  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ , and  $\theta_{m+1}$ , are needed to fully describe the behavior of the monitoring trace in the present layer. Because the thickness  $d = r \times t$ , where  $r$  is the deposition rate and  $t$  is the deposition time, a relationship between  $T$  and  $r$  or  $t$  is also established. If the deposition rate is constant, the accumulated phase  $\theta_{m+1}$  is proportional to the deposition rate. Therefore, during the deposition process, the deposition rate may be retrieved with high accuracy by fitting the in-situ measurement of the monitoring trace to the four parameters.

20

With this groundwork, the process of Fig. 6 will now be described in greater detail. However, prior to initiating the process of Fig. 6, a design for the SWP filter is prepared. The process of designing a SWP edge filter is very similar to that of LWP filter, with several exceptions. First, the initial structure is  $(0.5L \text{ H } 0.5L)^N$ . Second, the shorter wavelength edge of the QW stopband should be aligned with the desired cutoff wavelength. Third, if the first layer next to the substrate is a low index layer, it should be removed. Finally, it is advantageous, but not necessary, to artificially increase the optical thickness of the first layer.

With the design of the SWP filter at hand, the design data, as well as deposition rate data are received as input at 601. The input data has the content and format described with reference to 501 in Fig. 5. At 602, the transmission curves  $T_i$  vs.  $d$  at a series of wavelengths for each  $i^{\text{th}}$  layer are calculated as described with reference to 502 in Fig. 5. Choosing the best monitoring wavelength  $\lambda_m$  at 603 is similar to that described with reference to 503 in Fig. 5, except that it may be more advantageous to select  $\lambda_m$  by taking an average of a series of wavelengths than it is for an LWP filter.

At 604, the  $\beta$  parameters are calculated for each layer at wavelength  $\lambda_m$  using equations (1) through (8) described above. At 605, a determination is made as to which layers should have their deposition duration controlled by optical monitoring and which layers should have their deposition duration timed using an expected deposition time  $t_i$ . Such determination is made by simulating the deposition of each layer and selecting the layers having the least amount of simulated error to be optically monitored. Deposition of the other layers will be timed using the expected deposition time  $t_i$ .

The simulation process occurs by executing the processes of 606 to 617 as described below. However, actual deposition does not occur at 606 and 607, the processing described at 609 is skipped, and instead of measuring  $T_m$  vs.  $t$  at 612, it is generated.  $T_m$  vs.  $t$  is generated by adding random noise to the theoretical data  $T_i$  vs.  $d$  at  $\lambda_m$  from 602 and 603. In the exemplary embodiment, 0.2% peak-to-peak random noise is used, and the maximum amount of error ("threshold") to select a layer for optical monitoring is to have no more than about 0.5% error from the theoretical resulting thickness  $d_i$ . The error calculation, in this regard, is described in more detail below with reference to 613. The layers that are simulated to exceed the threshold

amount of error are flagged to have their deposition duration controlled by the best estimate of the deposition rate  $r_i$  for that layer or from an average of the rates of the previous layers of like material (typically 10 to 20 such layers).

Having determined which layers are to be optically monitored at 605, deposition begins  
5 at 606. In particular, a substrate is loaded into the deposition apparatus 400, the apparatus 400 is pumped down to vacuum, and deposition of the first layer (current layer  $i$ ) is initiated at 607. At 608, the expected deposition time  $t_i$  for layer  $i$  is calculated as the desired thickness  $d_i$  divided by the estimated deposition rate for the layer  $r_i$  or from an average of the rates of the previous layers of like material. It should be noted, however, that calculation of the expected deposition time  $t_i$   
10 at 608 may be calculated prior to beginning actual deposition of the current layer  $i$  at 607.

At 609 it is determined whether the current layer  $i$  was identified as an optically monitored layer at 605. If not, deposition occurs until the expected deposition time  $t_i$  expires, and the deposition apparatus is configured for deposition of the next layer, as shown at 610, 611, and 616. If it is determined that the current layer  $i$  is an optically monitored layer at 609, the  
15 transmission  $T_m$  of the current layer is measured at 612 as a function of actual time  $t$  transpired, thereby producing a measured curve  $T_m$  vs.  $t$ , until about 95% of time  $t_i$  has elapsed. Once about 95% of the time  $t_i$  has elapsed, a new layer time is calculated at 613 to 615.

To elaborate, at 613,  $T_i$  vs.  $t$  (where  $t = d/r$  and  $T_i$  vs.  $d$  at  $\lambda_m$  was calculated and selected at 602 and 603, respectively) is fit to the measured curve  $T_m$  vs.  $t$  from 612. The  $T_i$  vs.  $t$  curve is  
20 fit to the measured curve  $T_m$  vs.  $t$  by using a function that minimizes the error between the two curves by varying  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$ , and the deposition rate  $r$ . An example of such a function is the *Levenberg-Marquardt* method implemented under the name “mrqmin()” in the book Numerical Recipes in C: The Art of Scientific Computing, by Press, W.H.; Teukolsky, S.A.; Vetterling, W.T.; and Flannery, B.P., 2nd ed., Cambridge University Press, Cambridge, 1995. These  
25 calculations result in a calculated deposition rate for the current layer  $i$ , or  $r_{ci}$ .

Because of noise in the overall system and the sensitivity of the high-performance SWP filters to small layer-thickness errors, the calculated rates  $r_{ci}$  tend to be insufficiently accurate if used directly. Therefore, a “best rate” for each layer,  $r_{bi}$ , is calculated at 614 as the average of the calculated rate  $r_{ci}$  and calculated rates  $r_{cj}$ ,  $r_{c,j+2}$ ,  $\dots$ ,  $r_{c,i-4}$ ,  $r_{c,i-2}$  for a certain number  $(i - j)/2$  of

previous layers of the same material. In other words, the best rate is a rolling average of the current and previous rates from layers of like material within a certain window. Typically, this window includes about 20 layers.

At 615, the best rate  $r_{bi}$  is used to calculate the layer termination time  $t_i = d/r_{bi}$ , and the layer is terminated when the clock reaches this time. Once deposition of the current layer is complete, the deposition apparatus 400 of Fig. 4 is reconfigured at 616 to start depositing the material associated with the next layer  $i + 1$ , and the process loops back to 607. However, if all layers have been deposited, manufacturing of the SWP filter is complete, and the process ends at 617.

Although the exemplary methods have been described in the context of manufacturing optical edge filters, those skilled in the art will appreciate that such methods apply to manufacturing other types of optical filters, or may easily be modified to manufacture other types of optical filters. For instance, the methods are useful for any filter having even a single layer of critical thickness, such as a Fabry-Perot interferometer, where the layer of critical thickness is the resonance cavity. Further, although the exemplary methods are disclosed as having a particular sequence of events, one skilled in the art will appreciate that many of these events may occur in a different order without departing from the scope of the invention. Accordingly, the methods of the present invention are not limited to producing optical edge filters and are not limited to the particular ordering of events described.

### **C. The Improved Filters**

The improved filters have the general structure schematically illustrated in Fig. 3 but are made by the processes described herein in connection with Figs. 5 and 6. In essence, the filters comprise a transparent substrate advantageously having a pair of optically flat planar major surfaces and a plurality of alternating layers of relatively high index and relatively low index materials. The materials are hard oxide coating materials such as  $\text{SiO}_2$ ,  $\text{Ta}_2\text{O}_5$ ,  $\text{Nb}_2\text{O}_5$ ,  $\text{HfO}_2$ ,  $\text{TiO}_2$ , and  $\text{Al}_2\text{O}_3$ .

The filters differ from conventional edge filters primarily in the number of layers (typically more than 100) and in that the thicknesses of layers deposited are controlled by the

processes of Figs. 5 and 6 to produce an edge steepness less than about 0.8%. The edge steepness, in this regard, is measured by dividing (a) the edge width from the 50% transmission wavelength to the optical density 6 (“OD6”) wavelength by (b) the 50% transmission wavelength. Accordingly, lower steepness values indicate greater slope. Optical density (“OD”) is a measure of the blockage encountered by impinging light and is defined as follows:

$$OD = -\log_{10}(T) \quad (9)$$

where T is the transmittance having a value between zero and one. OD6 therefore corresponds to a transmittance of  $10^{-6}$ .

Advantageously the edge steepness, as defined above, is less than about 0.463%. However, the edge steepness may be increased if necessary. In addition, the filters so made exhibit an average transmission above about 93% and preferably above at least 95% with ripple below about 2% in the operating range. In the case of a filter for a 532 nanometer laser line, the operating range is approximately 200 nm from the wavelength at which the filter reaches full transmission. One skilled in the art will appreciate that the operating range is larger with filters for laser lines at higher wavelengths. Transmission of the filters does not drop below 93% and preferably 95% in the operating range. The filters thus provide performance exceeding that of the highest performing conventional soft-coating filters with a more robust and durable hard-coated structure.

The filters can be constructed with an edge wavelength as short as 325 nm and as long as 1064 nm. It is contemplated that this range may be extended from as low as 250 nm to as high as 2,000 nm using the materials listed above.

The invention can now be more clearly understood by consideration of the following specific examples.

#### **Example 1: Long-Wave-Pass Edge Filter**

A steep-edge LWP filter for a 532 nanometer laser line was designed and fabricated in accordance with the LWP design strategy described above. The filter is to block the 532 nanometer laser light and light of lesser wavelength but to transmit light of longer wavelength.

Appendix A gives the layer structure of the design. The optical thickness is given in units of quarter wavelengths (QWs) at the monitoring wavelength of 568 nanometers. The layers are counted from the substrate outward toward air. The substrate is BK7 glass, marketed by vendors such as Schott Glass. The design has 180 layers with a total metric thickness of 12.7 micrometers.

Fig. 7 shows the theoretical and measured transmission spectra of the resulting 532 nm LWP edge filter in accordance with the invention. Curve 701 is the theoretical spectrum, curve 702 is the measured spectrum and curve 703 is the laser wavelength line at 532 nm.

Fig. 8 illustrates the optical density spectrum of the 532 nm LWP filter. As previously discussed, optical density (OD) is a measure of the blockage encountered by impinging light. Curve 801 is the theoretical OD spectrum, curve 802 is the measured OD spectrum, and curve 803 is the 532 nm laser line. It is noteworthy that the filter edge is so steep that it is difficult to measure with standard spectrophotometry equipment. The apparent “kink” between OD3 and 4 is not real, but rather is produced by the measuring instrument.

### **Example 2: Short-Wave-Pass Edge Filter**

A steep-edge SWP filter was designed and fabricated in accordance with the SWP design strategy described above. The filter is to block the 532 nanometer laser light and light of greater wavelength but to transmit light of shorter wavelength. Appendix B provides the layer structure of the design. The optical thickness is given in QWs at the monitoring wavelength of 510 nm. The substrate is BK7 glass. The design has 180 layers with a total metric thickness of 15.1 micrometers.

Fig. 9 illustrates the theoretical and measured transmission spectra 901 and 902, respectively, of the realized SWP filter. The laser line at 532 nm is shown at 903.

Fig. 10 shows the theoretical and measured optical density spectra 1001 and 1002, respectively, and the laser line 1003. The apparent “kink” that occurs between about OD 4 and 5 is due to the limitations of the measuring instrument, not the filter.

### **Comparison of the Inventive Filters With Conventional Filters**

As compared with conventional edge filters, the filters of the invention provide substantially higher edge steepness and passband transmission. Table 1 below presents a comparison between the inventive filters and the best known examples of conventional soft-coated and hard-coated thin-film edge filters. The data in Table 1 are from filters designed to block a 532 nm laser, where the 50% transmission wavelength is 540 nm. Accordingly, steepness (in percentage) is calculated as Edge Width divided by 540 nm.

**Table 1**

|                             | <b><u>Hard-Coated</u></b> | <b><u>Soft-Coated</u></b>   | <b><u>Inventive</u></b> |
|-----------------------------|---------------------------|-----------------------------|-------------------------|
| Edge Width (50% to OD4)     | $\approx 3 \text{ nm}^*$  | $\approx 4 \text{ nm}^*$    | $< 1.4 \text{ nm}$      |
| Edge Steepness (50% to OD4) | $\approx 0.556\%$         | $\approx 0.741\%$           | $< 0.259\%$             |
| Edge Width (50% to OD6)     | $\approx 5 \text{ nm}^*$  | $\approx 6 \text{ nm}^{**}$ | $< 2.5 \text{ nm}$      |
| Edge Steepness (50% to OD6) | $\approx 0.926\%$         | $\approx 1.111\%$           | $< 0.463\%$             |
| Laser-Line Attenuation (OD) | $> 6.0$                   | $> 5.0$                     | $> 6.0$                 |
| Average Transmission        | $> 90\%$                  | $> 85\%$                    | $> 95\%$                |

\* Estimated specification

\*\* Edge width for 50% to OD5 (necessarily less than Edge Width for 50% to OD6)

As can be seen from the table, the filters of the present invention provide an improvement in edge steepness and a reduction of optical loss in the transmission band (optical loss defined as 100% - Average Transmission).

In addition, the inventive filters provide a substantial improvement in reliability and durability over soft-coated thin-film filters (the highest performing conventional filters) because the inventive filters are made exclusively with hard oxide glass, whereas soft-coated filters

contain soft salts and organic materials susceptible to damage by humidity and temperature extremes.

#### **D. Applications of the Filters**

5 The inventive filters can improve the performance of a variety of optical analysis systems that excite a sample of material with light of a first wavelength to produce a measurable or viewable response of light at a second wavelength different from the first. Such systems, which include Raman spectroscopy and fluorescence microscopy, are typically of the construction schematically shown in Fig. 2. They comprise a source of excitation light, an optical path coupling the excitation light to the sample, and a first filter in the path between the source and  
10 the sample for blocking light at some wavelengths different from the excitation light. They also include an optical path coupling the optical response light from the sample to an analyzer or viewer, and a second filter in the path between the sample and the analyzer or viewer for blocking some light other than the optical response. As previously discussed, the first and/or second filters can be long-wave-pass and/or short-wave-pass optical edge filters. It is also  
15 possible that one of the filters, *e.g.* the second, can be a notch filter.

One or more optical edge filters in accordance with the invention can substantially improve the performance of such optical analytical systems. The greater steepness of edge slope permits return of response wavelengths closer to the excitation wavelength providing an increase in the information content of the returned response. The reduction in transmission loss means  
20 that the enhanced information return response will be at higher brightness, and the greater hardness and durability of the inventive filters permits a more robust and versatile instrument.

It is to be understood that the exemplary embodiment is merely illustrative of the present invention and that many variations of the above-described embodiment can be devised by one skilled in the art without departing from the scope of the invention. It is therefore intended that  
25 all such variations be included within the scope of the following claims and their equivalents.